Nolan Stevenson, 109841202
ASEN 6055: HW2

Problem 1

1. Maximum Likelihood Estimator (MLE)

- Show the derivation of the maximum-likelihood estimator (MLE), starting from the log-likelihood function. Compute your car location by using MLE, and graphically show the location on the map.
- Change the GPS location observation errors to be uncorrelated, i.e., $\mathbf{R}^a = \begin{bmatrix} 0.6 & 0 \\ 0 & 1.2 \end{bmatrix}$ and $\mathbf{R}^b = \begin{bmatrix} 0.6 & 0 \\ 0 & 1.2 \end{bmatrix}$, and recompute the MLE of your car location.

  Discuss the difference from the previous case.

Starting w/ likelihood for a normal distro

$$[y|x] = \frac{1}{(2\pi)^{\frac{p}{2}} (\det(R))^{\frac{1}{2}}} e^{-\frac{1}{2}(y-Hx)^T R^{-1}(y-Hx)}$$

Log likelihood

$$\log\left(\frac{1}{(2\pi)^{\frac{p}{2}} (\det(R))^{\frac{1}{2}}}\right) + \log\left(e^{-\frac{1}{2}(y-Hx)^T R^{-1}(y-Hx)}\right)$$

$$\ell(p) = \log\left(\frac{1}{(2\pi)^{\frac{p}{2}} (\det(R))^{\frac{1}{2}}}\right) - \frac{1}{2}(y-Hx)^T R^{-1}(y-Hx)$$

term is indep of x! Thus just this

$\max(\ell(p))$,

$$\therefore \hat{x}_{MLE} = \underset{x}{\operatorname{argmax}}\, \ell(x) = \boxed{\underset{x}{\operatorname{argmin}} \left(\frac{1}{2}(y-Hx)^T R^{-1}(y-Hx)\right)}$$

Thus, deriv $=0$,

$$\frac{\partial}{\partial x}\left(\frac{1}{2}(y-Hx)^T R^{-1}(y-Hx)\right) = 0$$

Using matrix differentiation rules

$$= -\left(\frac{1}{2} \cdot H^T \cdot R^{-1} \cdot (y - Hx) + \frac{1}{2} \cdot H^T (R^{-1})^T \cdot (y - Hx)\right)$$

or just $= H^T R^{-1}(y - Hx)$ as $(R^{-1})^T = R^{-1}$

Then

$$H^T R^{-1}(y - Hx) = 0$$

$$H^T R^{-1} y - H^T R^{-1} H x = 0$$

$$H^T R^{-1} H x = H^T R^{-1} y$$

$$\boxed{\hat{x}_{MLE} = (H^T R^{-1} H)^{-1} H^T R^{-1} y}$$

Now, solving problem, we know $H = I$ as $y \approx x$

Thus,

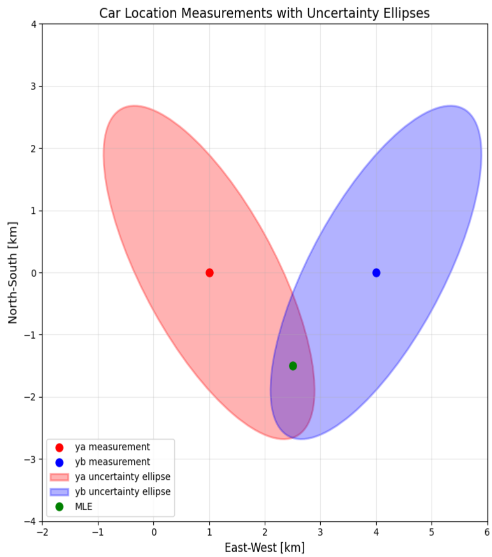
$$\boxed{\hat{x}_{MLE} = \left(R^{a^{-1}} + R^{b^{-1}}\right)^{-1}\left(R^{a^{-1}} y^a + R^{b^{-1}} y^b\right)}$$

summation

In python, this gives:

$$\boxed{\hat{x} = [2.5, -1.5]}$$

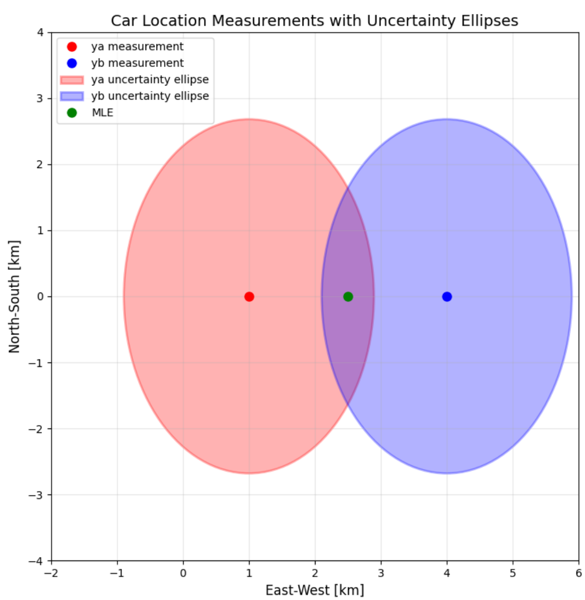
or graphically $\longrightarrow$



Now uncorrelated, $R^a = \begin{bmatrix} 0.6, 0 \\ 0, 1.2 \end{bmatrix}$ $R^b = \begin{bmatrix} 0.6, 0 \\ 0, 1.2 \end{bmatrix}$

$$\boxed{\hat{x} = [2.5, 0]}$$

The difference is now there is no North-south component in $\hat{x}$. The cars aren't skewed in that dimension which previously allowed north-south even w/o a meas in that direction.



Problem 2

2. Maximum A Posterior Estimator (MAP)

- Let's suppose that there is prior information about the car location given as

$$\mathbf{x} \sim \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1.2 & 0.6 \\ 0.6 & 0.6 \end{bmatrix}\right)$$

and compute your car location by using the MAP estimator. Use the original GPS location observation errors. Graphically show the location on the map, and discuss the difference from Problem 1 (first case) if any.

- Describe how the inference may change if the Bayesian approach is adopted.

prior $x = [0, 0]$, $P = \begin{bmatrix} 1.2, & 0.6 \\ 0.6, & 0.6 \end{bmatrix}$

w/ $y^a = [1, 0]$ $R^a = \begin{bmatrix} 0.6, & -0.6 \\ -0.6, & 1.2 \end{bmatrix}$ $y^b = [4, 0]$ $R^b = \begin{bmatrix} 0.6 & 0.6 \\ 0.6 & 1.2 \end{bmatrix}$

Where $\hat{x}_{MAP} = x^b + BH^T(HBH^T + R)^{-1}(y - Hx^b)$

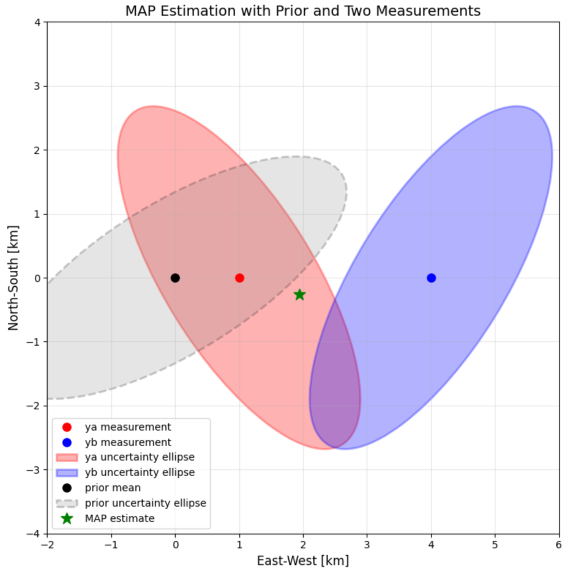
where we will stack meas and R such that

$$y = \begin{bmatrix} 1 \\ 0 \\ 4 \\ 0 \end{bmatrix} \quad R = \begin{bmatrix} 0.6 & -0.6 & 0 & 0 \\ -0.6 & 1.2 & 0 & 0 \\ 0 & 0 & 0.6 & 0.6 \\ 0 & 0 & 0.6 & 1.2 \end{bmatrix} \quad H = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$$

needed for matrix math

Result is

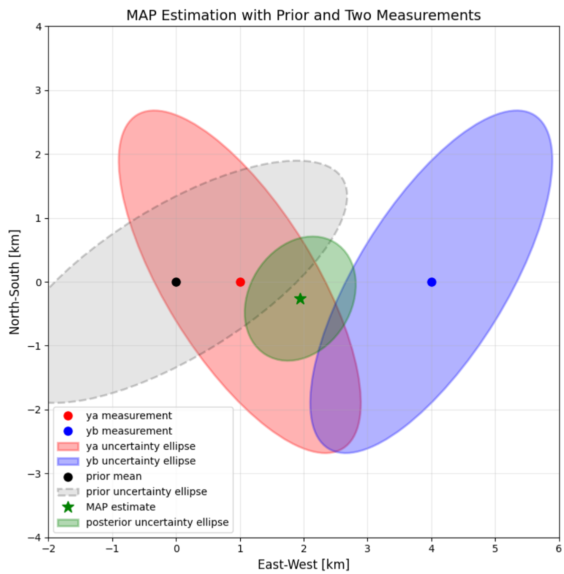
$$\boxed{\hat{x}_{MAP} = [1.947, -0.263]}$$



This approach differs from MLE by using the prior x and P.
This causes the estimate to go left, compared to before w/ a in the middle estimate.

The bayesian approach differs by also getting the P_{post}.
Which is the main advantage of Bayesian approach, you can get the cov.

I did the Bayesian using information update and get a cov:



But, same estimate!

Problem 3

3. Suppose that the car is moving in the south-east direction at the speed of 7.2 km/hr because it is being towed. Track the car location by using the Kalman filter. Use the prior car location information given in Problem 2, and the original GPS location observation errors. Assume that errors associated with a dynamical model for the motion of car (process noises) $\boldsymbol{\epsilon}^{\mathrm{x}}$ are distributed according to

$$\boldsymbol{\epsilon}^{\mathrm{x}} \sim \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0.04 & 0 \\ 0 & 0.04 \end{bmatrix}\right)$$

Two GPS devices report the following locations every 1 minute:

$$[\mathbf{y}_0^a \;\; \mathbf{y}_1^a \;\; \mathbf{y}_2^a \;\; \mathbf{y}_3^a \;\; \mathbf{y}_4^a \;\; \mathbf{y}_5^a \;\; \mathbf{y}_6^a] = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & -0.1 & -0.2 & -0.3 & -0.4 & -0.5 & -0.5 \end{bmatrix}$$

$$[\mathbf{y}_0^b \;\; \mathbf{y}_1^b \;\; \mathbf{y}_2^b \;\; \mathbf{y}_3^b \;\; \mathbf{y}_4^b \;\; \mathbf{y}_5^b \;\; \mathbf{y}_6^b] = \begin{bmatrix} 4 & 4 & 4 & 4 & 4 & 4 & 4 \\ 0 & -0.1 & -0.2 & -0.3 & -0.4 & -0.5 & -0.5 \end{bmatrix}$$

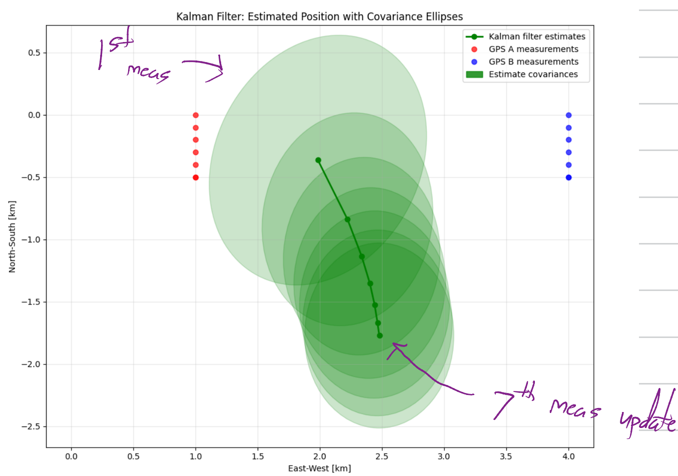
Assume using a constant position KF (as Q only given for pos)

w/ $x_0 = [0, 0]$ and standard linear KF update eqs with

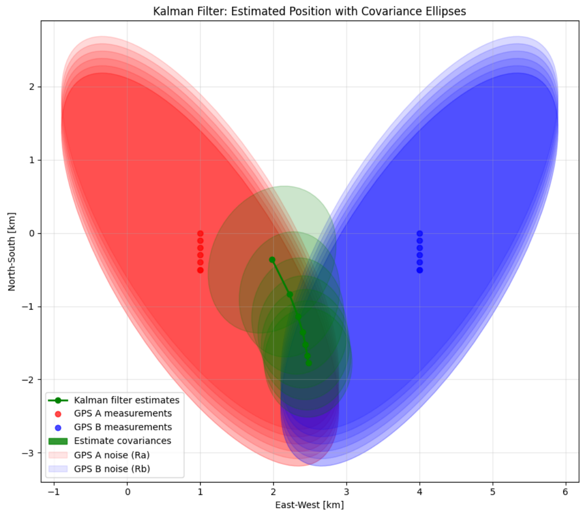
$$P_0 = \begin{bmatrix} 1.2, & 0.6 \\ 0.6, & 0.6 \end{bmatrix} \quad R^a = \begin{bmatrix} 0.6 & -0.6 \\ -0.6 & 1.2 \end{bmatrix} \quad R^b = \begin{bmatrix} 0.6 & 0.6 \\ 0.6 & 1.2 \end{bmatrix}$$

and $H = I_{2\times2}$, $F = I_{2\times2}$

Iteratively updating w/ meas results in:



Result does look a bit wild, but given the R's on the measurements, it does make sense according to their covs →



What is the prob of finding car @ (2, −2) @ 6 mins?

Use final state and cov:

$$x_6 = [2.4785, -1.7716] \qquad P_6 = \begin{bmatrix} 0.0601, & 8e{-}5 \\ 8e{-}5 & 0.093 \end{bmatrix}$$

Probability of multivariate Gaussian:

PDF

$$p(x|\mu) = (2\pi)^{-\frac{k}{2}} \det(\Sigma)^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)\right)$$

Thus,

$$= \frac{1}{2\pi\sqrt{|P_6|}} e^{\left(-\frac{1}{2}\left(\begin{bmatrix}2\\-2\end{bmatrix}-x_6\right)^T P_6^{-1}\left(\begin{bmatrix}2\\-2\end{bmatrix}-x_6\right)\right)}$$

$$= \boxed{24.026\%}$$ for (2, −2) @ 6 mins

Problem 4

4. Show the equivalence of the Bayesian approach and variational approach to the univariate normal state estimation problem by equations.

Using $x \sim N(\mu_0, \sigma_0^2)$, compute Bayesian posterior and variational w/ minimized KL

Bayes	Variational
$p(x \mid y) = N(\mu_{post}, \sigma^2_{post})$	$KL(q \mid p(y)) = E[\log q(x)] - E[\log p(x)] - E[\log p(x \mid y)]$

For single scalar,

$$\frac{1}{\sigma^2_{post}} = \frac{1}{\sigma_0^2} + \frac{1}{R}$$

and $$\frac{\mu_{post}}{\sigma^2_{post}} = \frac{\mu_0}{\sigma_0^2} + \frac{y}{R}$$

Giving:

$$\boxed{\mu_{post} = \sigma^2_{post}\left(\frac{\mu_0}{\sigma_0^2} + \frac{y}{R}\right) \qquad \sigma^2_{post} = \left(\frac{1}{\sigma_0^2} + \frac{1}{R}\right)^{-1}}$$

Variational:

$$\mathcal{L}(m, s^2) = \frac{1}{2}\log(s^2) + \frac{(m-\mu_0)^2 + s^2}{2\sigma_0^2} + \frac{(y-m)^2 + s^2}{2R}$$

$$\frac{\partial}{\partial m} \rightarrow \frac{m-\mu_0}{\sigma_0^2} + \frac{m-y}{R}$$

$$m\left(\frac{1}{\sigma_0^2} + \frac{1}{R}\right) = \frac{\mu_0}{\sigma_0^2} + \frac{y}{R}$$

Thus, set $=0$,

$$\boxed{m = \left(\frac{\mu_0}{\sigma_0^2} + \frac{y}{R}\right)\left(\frac{1}{\sigma_0^2} + \frac{1}{R}\right)^{-1}}$$

$$\frac{\partial}{\partial s^2} \rightarrow \frac{1}{s^2} + \frac{1}{\sigma_0^2} + \frac{1}{R}$$

Thus,

$$\boxed{s^2 = \left(\frac{1}{\sigma_0^2} + \frac{1}{R}\right)^{-1}}$$

$= s^2$

so, $$\boxed{m = s^2\left(\frac{1}{\sigma_0^2} + \frac{1}{R}\right)^{-1}}$$

Thus, we have

Bayes — Variational

$$\mu_{post} = \sigma_{post}^2 \left(\frac{\mu_0}{\sigma_0^2} + \frac{1}{R} \right) \approx m = s^2 \left(\frac{\mu_0}{\sigma_0^2} + \frac{1}{R} \right)$$

and

$$\sigma_{post}^2 = \left(\frac{1}{\sigma_0^2} + \frac{1}{R} \right)^{-1} \approx s^2 = \left(\frac{1}{\sigma_0^2} + \frac{1}{R} \right)^{-1}$$

where $\mu_{post} = m$, $\sigma_{post}^2 = s^2$